

nucleases of Pol I-type DNA polymerases. These are the FEN1 (Flap EndoNuclease), RAD2, and XPG (Xeroderma Pigmentosa-complementation group G) proteins. These proteins are involved in DNA repair, and have been shown to favor the cleavage of structures that resemble a 5' arm that has been displaced by an extending primer during polymerization, similar to the model depicted in Figure 16B. Similar DNA repair enzymes have been isolated from single cell and higher eukaryotes and from archaea, and there are related DNA repair proteins in eubacteria. Similar 5' nucleases have also be associated with bacteriophage such as T5 and T7.

Recently, the 3-dimensional structures of DNAP_{Taq} and T5 phage 5'-exonuclease (Figure 69) were determined by X-ray diffraction [Kim *et al.* (1995) Nature 376:612 and Ceska *et al.* (1995) Nature 382:90). The two enzymes have very similar 3-dimensional structures despite limited amino acid sequence similarity. The most striking feature of the T5 5'-exonuclease structure is the existence of a triangular hole formed by the active site of the protein and two alpha helices (Figure 69). This same region of DNAP_{Taq} is disordered in the crystal structure, indicating that this region is flexible, and thus is not shown in the published 3-dimensional structure. However, the 5' nuclease domain of DNAP_{Taq} is likely to have the same structure, based its overall 3-dimensional similarity to T5 5'-exonuclease, and that the amino acids in the disordered region of the DNAP_{Taq} protein are those associated with alpha helix formation. The existence of such a hole or groove in the 5' nuclease domain of DNAP_{Taq} was predicted based on its substrate specificity [Lyamichev *et al.*, *supra*].

It has been suggested that the 5' arm of a cleavage structure must thread through the helical arch described above to position said structure correctly for cleavage (Ceska *et al.*, *supra*). One of the modifications of 5' nucleases described herein opened up the helical arch portion of the protein to allow improved cleavage of structures that cut poorly or not at all (*e.g.*, structures on circular DNA targets that would preclude such threading of a 5' arm). The gene construct that was chosen as a model to test this approach was the one called Cleavase® BN, which was derived from DNAP_{Taq} but does not contain the polymerase domainn (Ex. 2). It comprises the

entire 5' nuclease domain of DNAP Taq, and thus should be very close in structure to the T5 5' exonuclease. This 5' nuclease was chosen to demonstrate the principle of such a physical modification on proteins of this type. The arch-opening modification of the present invention is not intended to be limited to the 5' nuclease domains of DNA polymerases, and is contemplated for use on any structure-specific nuclease which includes such an aperture as a limitation on cleavage activity.

The opening of the helical arch was accomplished by insertion of a protease site in the arch. This allowed post-translational digestion of the expressed protein with the appropriate protease to open the arch at its apex. Proteases of this type recognize short stretches of specific amino acid sequence. Such proteases include thrombin and factor Xa. Cleavage of a protein with such a protease depends on both the presence of that site in the amino acid sequence of the protein and the accessibility of that site on the folded intact protein. Even with a crystal structure it can be difficult to predict the susceptibility of any particular region of a protein to protease cleavage. Absent a crystal structure it must be determined empirically.

In selecting a protease for a site-specific cleavage of a protein that has been modified to contain a protease cleavage site, a first step is to test the unmodified protein for cleavage at alternative sites. For example, DNAPTaq and Cleavase® BN nuclease were both incubated under protease cleavage conditions with factor Xa and thrombin proteases. Both nuclease proteins were cut with factor Xa within the 5' nuclease domain, but neither nuclease was digested with large amounts of thrombin. Thus, thrombin was chosen for initial tests on opening the arch of the Cleavase® BN enzyme.

In the protease/Cleavase® modifications described herein the factor Xa protease cleaved strongly in an unacceptable position in the unmodified nuclease protein, in a region likely to compromise the activity of the end product. Other unmodified nucleases contemplated herein may not be sensitive to the factor Xa, but may be sensitive to thrombin or other such proteases. Alternatively, they may be sensitive to these or other such proteases at sites that are immaterial to the function of the nuclease sought to be modified. In approaching any protein for modification by addition of a

protease cleavage site, the unmodified protein should be tested with the proteases under consideration to determine which proteases give acceptable levels of cleavage in other regions.

Working with the cloned segment of DNAPTaq from which the Cleavase® BN protein is expressed, nucleotides encoding a thrombin cleavage site were introduced in-frame near the sequence encoding amino acid 90 of the nuclease gene. This position was determined to be at or near the apex of the helical arch by reference to both the 3-dimensional structure of DNAPTaq, and the structure of T5 5' exonuclease.

The encoded amino acid sequence, LVPRGS, was inserted into the apex of the helical arch by site-directed mutagenesis of the nuclease gene. The proline (P) in the thrombin cleavage site was positioned to replace a proline normally in this position in Cleavase® BN because proline is an alpha helix-breaking amino acid, and may be important for the 3-dimensional structure of this arch. This construct was expressed, purified and then digested with thrombin. The digested enzyme was tested for its ability to cleave a target nucleic acid, bacteriophage M13 genomic DNA, that does not provide free 5' ends to facilitate cleavage by the threading model.

While the helical arch in this nuclease was opened by protease cleavage, it is contemplated that a number of other techniques could be used to achieve the same end. For example, the nucleotide sequence could be rearranged such that, upon expression, the resulting protein would be configured so that the top of the helical arch (amino acid 90) would be at the amino terminus of the protein, the natural carboxyl and amino termini of the protein sequence would be joined, and the new carboxyl terminus would lie at natural amino acid 89. This approach has the benefit that no foreign sequences are introduced and the enzyme is a single amino acid chain, and thus may be more stable than the cleaved 5' nuclease. In the crystal structure of DNAPTaq, the amino and carboxyl termini of the 5'-exonuclease domain lie in close proximity to each other, which suggests that the ends may be directly joined without the use of a flexible linker peptide sequence as is sometimes necessary. Such a rearrangement of the gene, with subsequent cloning and expression could be accomplished by standard PCR recombination and cloning techniques known to those skilled in the art.